MODEL BUILDING AND LSTM-BASED SYSTEM IDENTIFICATION FOR IMPLANTABLE DEVICES

A Dissertation
Presented to
The Academic Faculty

By

Mi Zhou

In Partial Fulfillment
of the Requirements for the Degree
Master in the
School of Electrical and Computer Engineering

Georgia Institute of Technology

May 2020

MODEL BUILDING AND LSTM-BASED SYSTEM IDENTIFICATION FOR IMPLANTABLE DEVICES

Approved by:

Dr. Zhang, Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Ma
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. AlRegib
School of Electrical and Computer Engineering
Georgia Institute of Technology

Date Approved: April 22, 2020

A large part of mathematics which becomes useful developed with absolutely no desire to be useful, and in a situation where nobody could possibly know in what area it could become useful; and there were no general indications that it ever could be so.

John Von Neumann

This thesis is dedicated to my parents. Thank you for supporting me spiritually and financially.

ACKNOWLEDGEMENTS

Firstly, I wish to express my deepest gratitude to my parents, who always support me and encourage me to pursue higher education. Secondly, I owe great thanks to Dr. Zhang, for giving me numerous advice and instructions. I am also very grateful to all the committee members who read my thesis. Thirdly, I would like to express my special thanks to Ayca, who plays the role of both advisor and colleague during the whole process I did my thesis. Finally, thank you, this wonderful world.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Implantable medical devices (IMDs) have aroused a wide research interest because of its increased ability of monitoring and recording signals from human organs and tissues. There are numerous issues that are under researching, one of which is the thermal control problem. Human body has a certain tolerance to high temperature (such as fever). However, long-term accumulation of the thermal effect if the temperature is higher than the safe limit leads to detrimental effect. Thus efficient thermal management methods are of significant importance for IMDs. Before designing controllers, the first aim is to build a reliable model for the thermal effects of the IMDs.

In this thesis, three different system identification methods are explored for modeling thermal effects and their advantages and disadvantages are compared. First, a COMSOL model considering all the thermal effects of an IMD system is built. Then, a long short-term memory (LSTM) network is designed to predict the thermal dynamics of the IMDs both online and offline. For the validation of the LSTM algorithms, both COMSOL simulations and experiments are studied. The performance of the LSTM method is compared with that of a recursive predictor based subspace identification (RPBSID) method considering time complexity and prediction accuracy. The results based on COMSOL simulations indicate that the online LSTM algorithm outperforms the RPBSID algorithm in general except its higher computational cost. The offline LSTM algorithm has superiority for the time period when the convergence of the adaptive filters in the RPBSID algorithm is not yet achieved. Additionally, the results based on *in vitro* experiments show that both online LSTM and offline LSTM triumph over RPBSID based on the metric - best fit rate (BFR).

CHAPTER 1
INTRODUCTION

1.1 Implantable Medical Devices

An implantable medical device, as its name shows, is a device that is placed either partly or totally inside human bodies. Recently, most widely known IMDs are cardiometer, pacemaker, cochlear implants, replacement heart implants, and retinal prostheses, with which, medical care has been largely enhanced for patients. The purpose of the IMDs is to restore or maintain a normal body function, to detect some human body signals, such as heartbeat frequency, blood pressure and to provide electrical or mechanical assistance to human organs. Starting from the first pacemaker implant in 1958, numerous researches about IMDs have been developed, for which, the challenges lie on not only materials, size shrinkage, wireless communication, but also battery power, system delivery, biology rejection, patients' comfort level and so on [1]. Considering that the human body is a complex environment operated by mechanical, chemical and electrical mechanisms of a large amount of organs, tissues and cells, making an IMD commercialized and widely used is more difficult than theory. Take the implantable blood pressure measurement devices as an example: besides determining a location for installation that allows for better data processing and wireless communication in battery performance, we need also consider corrosion induced by chemicals in the blood and achieve thermal control to maintain a safe limit of the blood environment. As shown in [2, 3], a slight increase of temperature, such as 2 °C, can lead to a significant amount of damage to the human organs or physiological abnormalities such as angiogenesis or necrosis.

1.2 Motivation of the Research

Based on aforementioned requirements and potential risks, a lot of research has been conducted to analyze thermal effects of IMDs. Reference [4] presented the temperature increase of a dual-unit retinal prosthesis under different parameters, such as number of electrodes, the characteristics of the telemetry and different tissues. In [5], the thermal influence of the integrated 3-D Utah electrode array device implanted in the brain was investigated by both numerical simulations and experiments *in vitro* and *vivo* environment.

Most of above-mentioned researches focused on analyzing the temperature increases of human tissues under different current stimulation or human environments. The thermal models used are based on solving Pennes' bio-heat equation, which is not applicable in the real-time applications. Especially for the IMDs, real-time information is needed for the following-up controlling work. In this regard, it is of great significance to have a highly efficient algorithm for thermal modeling of IMDs.

Recently, [6] proposed an online modeling method to predict the thermal effect of IMDs. Additionally, a joint power and dynamic thermal management method for neural prosthetic IMDs was proposed and the results demonstrated that it can achieve 34% more operation time while satisfying the same thermal constraint [7]. Reference [8] applied a recursive subspace identification method to predict the thermal dynamics of IMDs. Both simulation and experimental results indicated the good performance of this method.

This research is to explore the feasibility of using neural network-based method (i.e., LSTM) to predict the thermal dynamics of bio-implants. The conventional system identification algorithm is used for comparison to verify the effectiveness and robustness of the LSTM-based approach.

1.3 Organization of the Thesis

This thesis is constructed with seven chapters. Chapter 1 introduces the present researches of IMDs and the existing problems in thermal modeling process, which then leads to our research objective. Chapter 2 is the background review of the advancement of the system identification algorithms. Then Chapter 3 presents the COMSOL model of the IMDs we are studying and the thermal effects it has. Subsequently, the algorithms of the LSTM and the RPBSID are presented in Chapter 4. In Chapter 5, simulation studies under eight different scenarios are presented to verify the effectiveness and robustness of the proposed algorithms. Two metrics are used for fair comparison of the algorithms. Similarly, Chapter 6 presents our experimental platform and the results of using the platform to compare above-mentioned algorithms. Finally, this thesis is concluded with a summary of contributions and a discussion of the future works in Chapter 7.

CHAPTER 2

BACKGROUND REVIEW

System identification has a wide usage in practical industrial productions, such as process control, aerospace, disk drives and embedded systems. For these applications, in consideration of model simplicity or short cycle time, a complicated physics-based model is hard to obtain and thus impractical. The main aim of system identification, as its name suggests, is to study the dynamic behavior of an existing structure by its measured output or input-output in discrete time signals. A basic system identification process can be presented as shown in Figure 2.1.

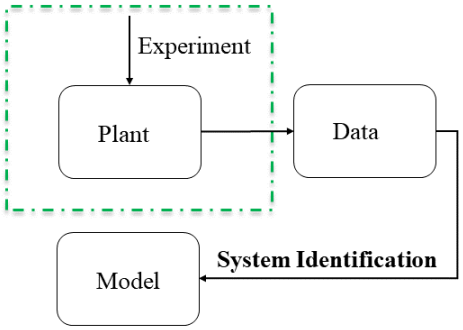


Figure 2.1: System identification process

In general, based on whether the model structure and parameters are known or not, the identification method can be classified as white-box identification, gray-box identification and black-box identification.

- *White-box identification*: estimate parameters of a physical model from data.

- *Gray-box identification*: estimate parameters from data while the model structure is known.
- *Black-box identification*: both model structure and parameters are unknown.

2.1 System Identification Methods: A review

2.1.1 Subspace Identification

There are many works and a strong theoretic basis for identifying linear time invariant (LTI) and linear parameter varying (LPV) systems, which are mainly divided into two categories: subspace identification (SID) and prediction error minimization (PEM). For a LTI system, some references combined subspace and optimization-based identification method by an iterative projected gradient search method to deal with the non-uniqueness problem [9]. Then, [10] extended this method to the identification of multi-variable nonlinear state space systems, such as bi-linear systems, composite local linear systems, Hammerstein systems and Wiener systems. Reference [11] proved that the l_2 loss function of one-step-ahead prediction for the LPV-MAX model has only one unique minimum under certain conditions on the input and scheduling signals. Even though SIMs and PEMs are widely used and perform well for linear systems, identification for nonlinear systems and large scale systems is sill facing a lot of difficulties.

2.1.2 Machine Learning-based Identification

In recent years, machine learning is universally studied in all kinds of fields due to the development of hardware and parallel computing. The machine learning-based system identification algorithm has been broadly studied because of its advantage in identifying systems without any prior knowledge of the systems' structure. From kernel-based non-parametric methods to deep neural networks, some papers focus on its application, while some researchers are interested in the robustness and theoretical explanation. As a trailblazer in

subspace system identification field, Lennart Ljung also made great contributions to general nonlinear black-box methods including neural networks and neuro-fuzzy modelling methods [12]. In book [13], Johan Suykens presented neural network-based nonlinear system identification methods, parametric-based methods, support vector machine and fuzzy models, which covers almost all the advanced black-box system identification techniques. Reference [14] provided both a theoretical and an experimental connection between system identification and machine learning. Moreover, deep neural network comes into view. Even though deep neural networks have already been universally introduced in computer vision and some other fields, there exists two main problems for applying deep learning methods in control theory, i.e., time complexity and robustness. For a practical control problem, achieving real-time control is extremely important; however, training a neural network is quite time-consuming. This problem of time complexity in real-time applications of deep learning methods can be solved in two ways. First, we can train the network with existing data offline, then perform the system identification process online with the obtained model. This approach is quick but doesn't perform well when facing new unknown data. The other approach is designing a simpler network structure and then repeating the train-predict-train cycle online with a small window of samples. This approach alleviates the time complexity problem and can satisfy the real-time requirement. Reference [15] presented an adaptive gradient learning method for recurrent neural networks to predict data series with anomalies and change points and realized online learning. In [16], the author proposed a new convex LSTM for the fast learning purpose both on system identification and control. Simulation and theoretical analysis verified that the proposed method outperforms the traditional LSTM in accuracy and speed. Some references focused on improving the structure of LSTM. For instance, [17] proposed a novel LSTM structure with a hierarchical recurrent networks and one multi-layer perception and designed a special learning algorithm based on back propagation. All in all, all these methods are trying to improve the performance in two aspects, online or offline.

2.2 System Identification Methods for Modeling the Thermal Effects of IMDs

The thermal effects are usually modelled as a partial differential equation (PDE) named Pennes bio-heat equation [7]. Despite the appearance of all kinds of system identification techniques, the references of thermal model identification is sparse. What's worse, in most of the cases, the model is approximated to a linear model. For example, [7] simplified the heating model to a linear model and then proposed an online system parameter estimation and control method for the implantable devices. Some papers exploited various numerical schemes combined with the differential transform and finite difference methods to analyze the temperature problem. For example, in [18], the authors presented a numerical study of the thermal increase in the human eye and head based on the numerical finite-difference time domain (FDTD) method. Reference [19] proposed a simplified one-dimensional bio-heat transfer model of the living tissues in the steady state based on the modified version of the Pennes bio-heat equation and derived its corresponding analytic periodic and non-periodic solutions. Reference [20] established a series solution for a 2D bio-heat Pennes conduction model with convective boundary by using classical Fourier method. Despite the high accuracy of these methods, the finite elements analysis (FEA)-based methods always need a lot of mathematical analysis and have high computation and time complexity, which makes it hard to be accomplished in the real control platform.

In view of the complexity of human body environment, regarding it as a black box and using system identification methods based on the inputs and outputs are of significant importance. Additionally, with advancements in GPU, the neural network (NN)-based system identification algorithm has already been explored for thermal effects. For example, [21] proposed a novel approach to automatically detecting the major heat source on a commercial multi-core microprocessor using an infrared thermal imaging setup by constructing LSTM networks. Reference [22] studied the use of NNs to develop models for nonlinear PDEs.

Based on the above reasons, we explored the LSTM-based system identification method for thermal dynamics modeling of IMDs. The conventional system identification algorithm is also realized for comparison. We believe, with the development of both theory and computing hardware, neural networks will be introduced in this field for practical applications in the near future.

CHAPTER 3

ANALYSIS AND SIMULATION OF IMD THERMAL DYNAMICS

Finite element analysis (FEA) has been widely used for thermal dynamics analysis of IMDs. In this chapter, a FEA model is implemented using the computational software COMSOL to capture the thermal dynamics of an IMD, which will also be used to validate the simplified thermal dynamic models later. Highlights of this chapter are two folds:

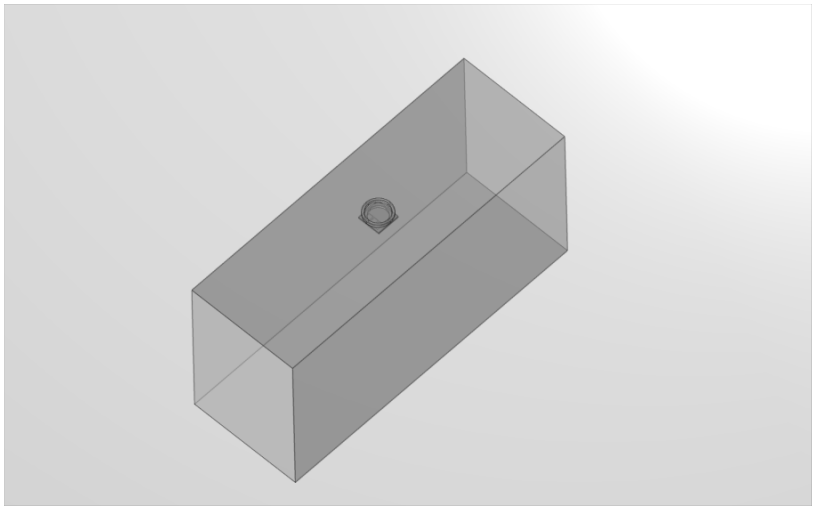
- A 3-D COMSOL model is built to capture the thermal dynamics of a temperature monitoring and management test vehicle (TMTV) we developed in a human environment. The geometry and materials properties of the model are set to approach the real human environment.
- Three thermal effects are analyzed to verify the effectiveness of the model.

3.1 COMSOL Model

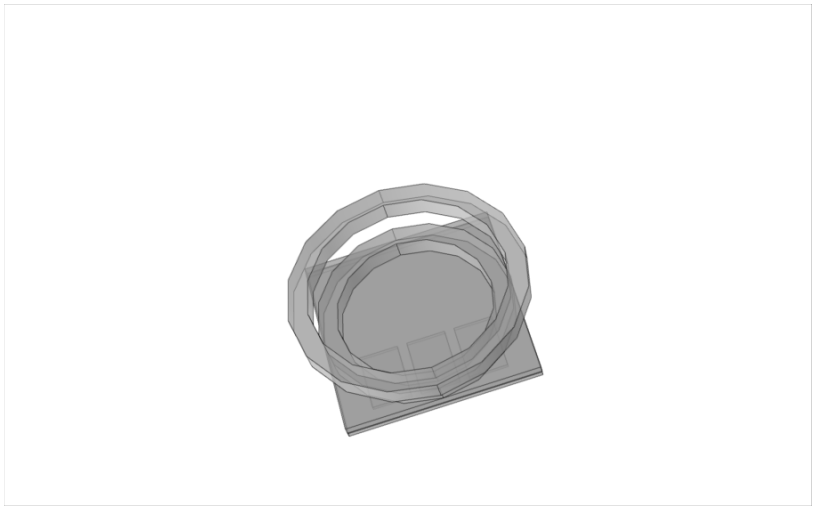
The COMSOL model is built with COMSOL 5.4 on Windows. As shown in Figure 3.1, a 3-D TMTV board (Figure 3.1 (b)) is put into the 3-D cube-like water-filled chamber (Figure 3.1 (a)) which is to emulate the human environment.

3.1.1 Geometry and Material Design

The simulated board is composed of two coils (namely, coil 1 and coil 2), two heat sources and a PCB board which has 6 temperature sensors on it. Coil 1 is placed outside of the chamber and coil 2 is attached on the board implanted inside the chamber. The coils are utilized for wireless power transfer and the board is the main carrier of our intelligent chip. There are two fluids at different locations of the model: one of them is water (initial temperature 37 °C) and the other is air (initial temperature 20 °C), which are located inside



(a)



(b)

Figure 3.1: (a) Simulated human environment, (b) Simulated TMTV board

and outside the chamber respectively. The materials used and their properties are shown in Table 3.1, where ”/” means there is no item in the cell and ”*” denotes the item in the cell is a function of the temperature (not a constant coefficient). Coil 1 has a radius of 50 mm while coil 2 has a radius of 40 mm. Both of the coils have a height of 10 mm. The sizes of the board, heat sources and chamber are shown in Table 3.2.

Table 3.1: PHYSICAL STRUCTURE OF SYSTEM

Name	**Material**	**Thermal conductivity [W/(m*K)]**	**Electrical conductivity [S/m]**
Coil 1	Copper	400	5.998e7
Coil 2	Copper	400	5.998e7
Board	FR4	0.3	0.004
Heat Sources	SiO2	1.4	0
Water	/	*	5.5e-3
Air	/	*	0

Table 3.2: SIZE OF GEOMETRY PART

Name	**Width [mm]**	**Depth [mm]**	**Height [mm]**
Board	100	100	8.5
Heat Sources	20	30	2
Chamber	1400	500	700

3.1.2 Illustration of the board

The PCB board contains two heat sources (H1 and H2) and six temperature sensors (S1 - S6). Figure 3.2 shows our target system, where the temperature readings at positions S1-S6 are the outputs and H1 and H2 are the inputs of this board. Thus, this is a multi-input and multi-output (MIMO) system. Temperature data obtained by the COMSOL software will be used to demonstrate the performance of the algorithms in Chapter 4.

3.2 Evaluation of the COMSOL Model

Although we cannot evaluate the accuracy of the COMSOL model using experiments now, analysis is conducted below to evaluate the soundness of the model from the following three thermal aspects [23]:

- Radiation: the transfer of energy by the emission of electromagnetic radiation, which is generated by the alternative current (AC) in coil 1 and coil 2 whose current is

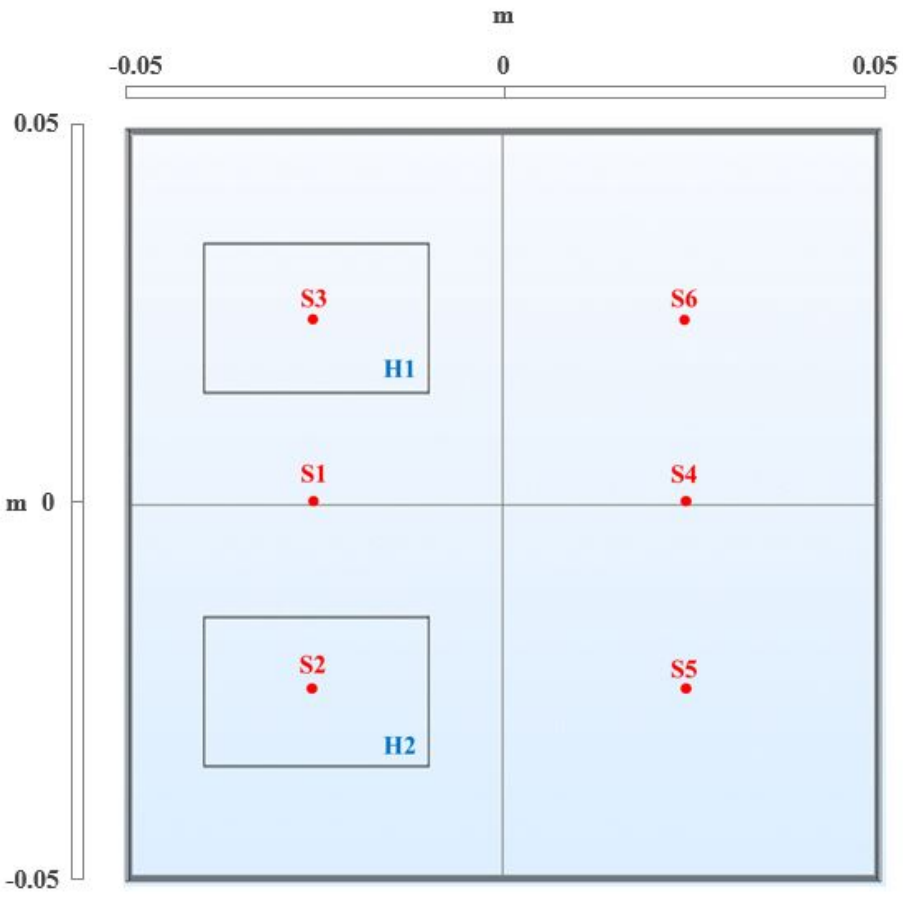


Figure 3.2: PCB Board implemented in COMSOL

excited due to induction.

- Convection: this is generated because heat produced by the heat sources moves to the water due to the fluid motion.
- Conduction: this kind of heat transfer exists within the PCB board, water and between the board and the water.

In the following subsections, we will introduce the models of all three kinds of thermal effects and use COMSOL simulation to demonstrate the soundness of the model.

3.2.1 Radiation

Radiation is an energy transfer via electromagnetic wave action. Since the energy is transported by electromagnetic radiation, no medium is needed. For the wireless model, coil 1 is passed through with AC current. The total emitted thermal energy is given by

$$E(T) = \epsilon \sigma T^4 \tag{3.1}$$

where $\sigma = 5.67 \times 10^{-8}$ $[\mathrm{W}/(\mathrm{m}^2\mathrm{K}^4)]$ is the Stefan-Boltzmann constant; T is the absolute temperature and ϵ is the thermal emissivity averaged over the thermal radiation spectrum. Due to the negligible radiative heat, in most cases, Equation 3.1 is not considered in the heat balance equation [23].

3.2.2 Convection

Convection is generated due to the thermal transfer from the board to the environment (water), which is modelled according to the Newton's law of cooling:

$$q_c = h_c(T_{board} - T_{water}) \tag{3.2}$$

where q_c is the heat flux (W^{-2}); h_c is the convection heat transfer coefficient ($\mathrm{WM}^{-2}\mathrm{K}^{-1}$) and T_{board} and T_{water} are the temperature of the board and the water respectively.

Since the board is coated and our research focuses on the temperature variations in the water, we don't consider the convection between coil 1 and the air and the convection between the board and the air. Only the convection between the board and the water is considered.

In our model, the water keeps flowing, which functions as cooling down the board and plays an important role in the thermal dynamics. To verify this conclusion, the input of the H1 is set as 0W and the input of the H2 is set as 10W. The only variable here is the velocity

of the water in the x direction, which is set as 0 m/s, 0.0022 m/s (average blood velocity) and 2 m/s respectively. Figure 3.3 shows the temperature at S1 under different velocities of the water. When the velocity of water is 0.0022 m/s, there is a temperature increase of 0.5 °C. Whilst when the velocity is 2 m/s, the temperature doesn't change, which implies the thermal equilibrium.

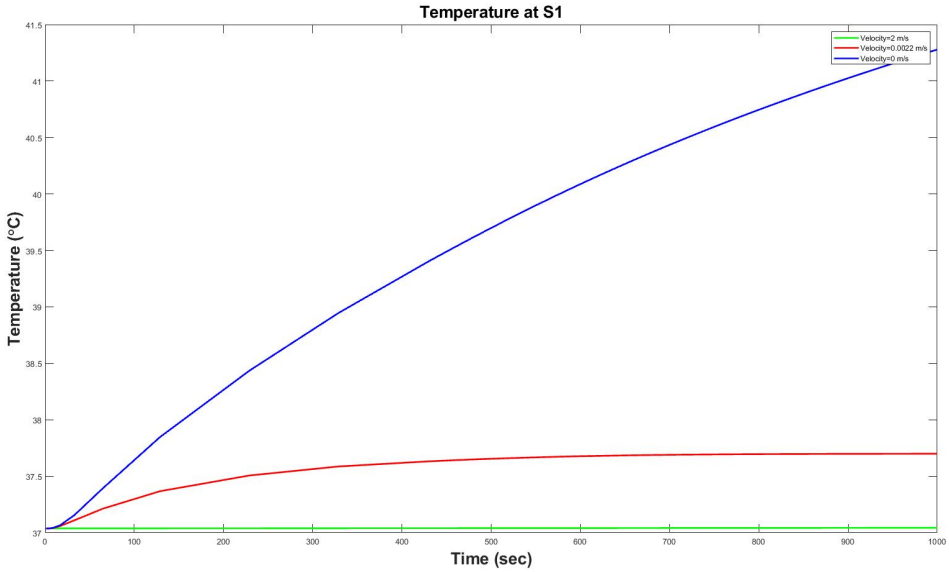


Figure 3.3: Temperature of S1 in the water with different velocities

3.2.3 Conduction

The conduction heating is governed by the Fourier law with equation

$$Q = -kA(T_2 - T_1)\Delta t/\Delta L \tag{3.3}$$

where $T_2 - T_1$ is the temperature difference; Δt is the length of time; A is the cross sectional area and $k = 0.63\ (\mathrm{Wm^{-1}K^{-1}})$ is the heat conductivity coefficient in water at 37 °C. Based on this, we can get the flow of heat (which is heat flux vector in COMSOL) due to thermal conduction as

$$f = -k(T_2 - T_1) \tag{3.4}$$

Since the highest temperature generally appears on the PCB board in our simulation, we mainly focus on the thermal effect induced by conduction between coil 2 and the board, and conduction between two heat sources and the board, both of which are generated by conduction in solid. By giving 1000 A current to the coil 2 and set the input power of heat sources as 0, we obtain Figure 3.4. As is shown in Figure 3.4, the board's temperature rises from 37 °C to more than 90 °C around the coil and the temperature decreases from coil 2 to the center of the board, which proves that the conduction between coil 2 and the board is working in our model.

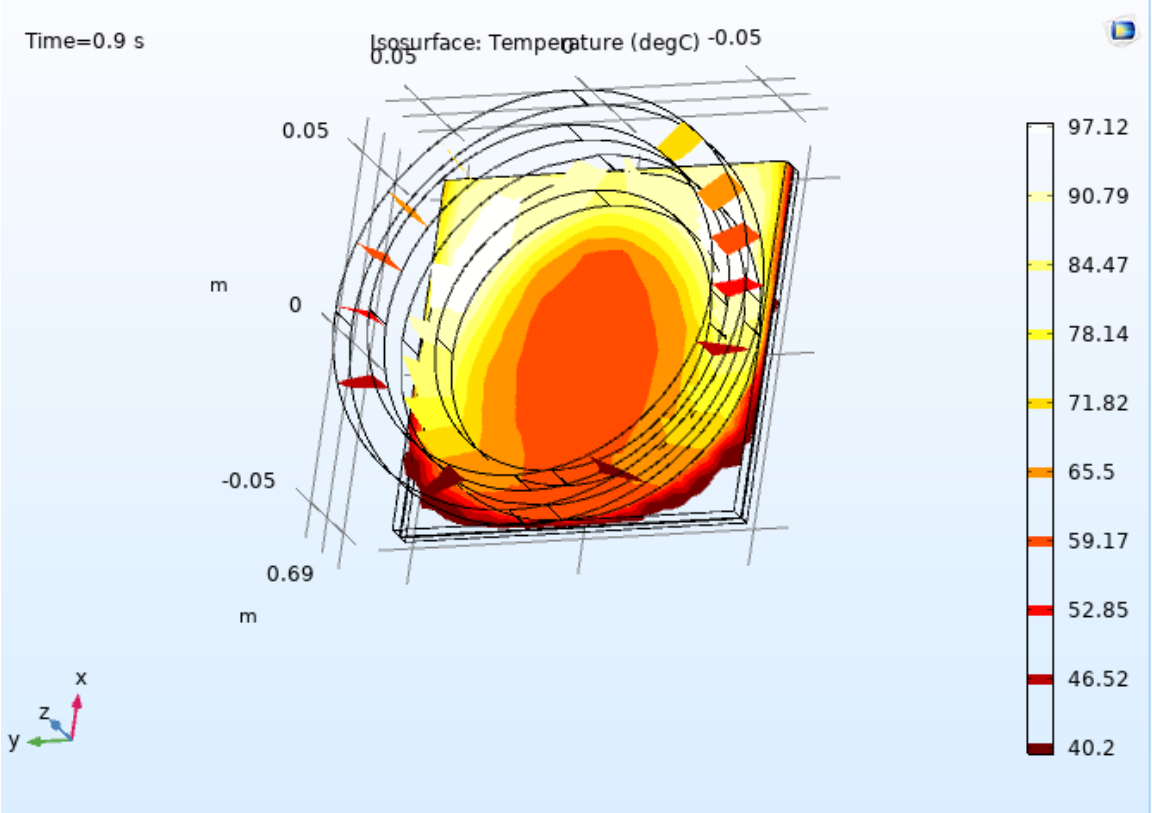


Figure 3.4: Conduction between coil 2 and the board

The conduction between heat sources and the board is shown in Figure 3.5 by providing a power of 10 W and 2 W to the two heat sources respectively, which aims to verify that the heat sources are working. Figure 3.5 shows that the temperature around two heat sources are higher than that in other positions. The heat source with a power input 10 W generates more heat than that of the other heat source. Please note that the values of these parameters

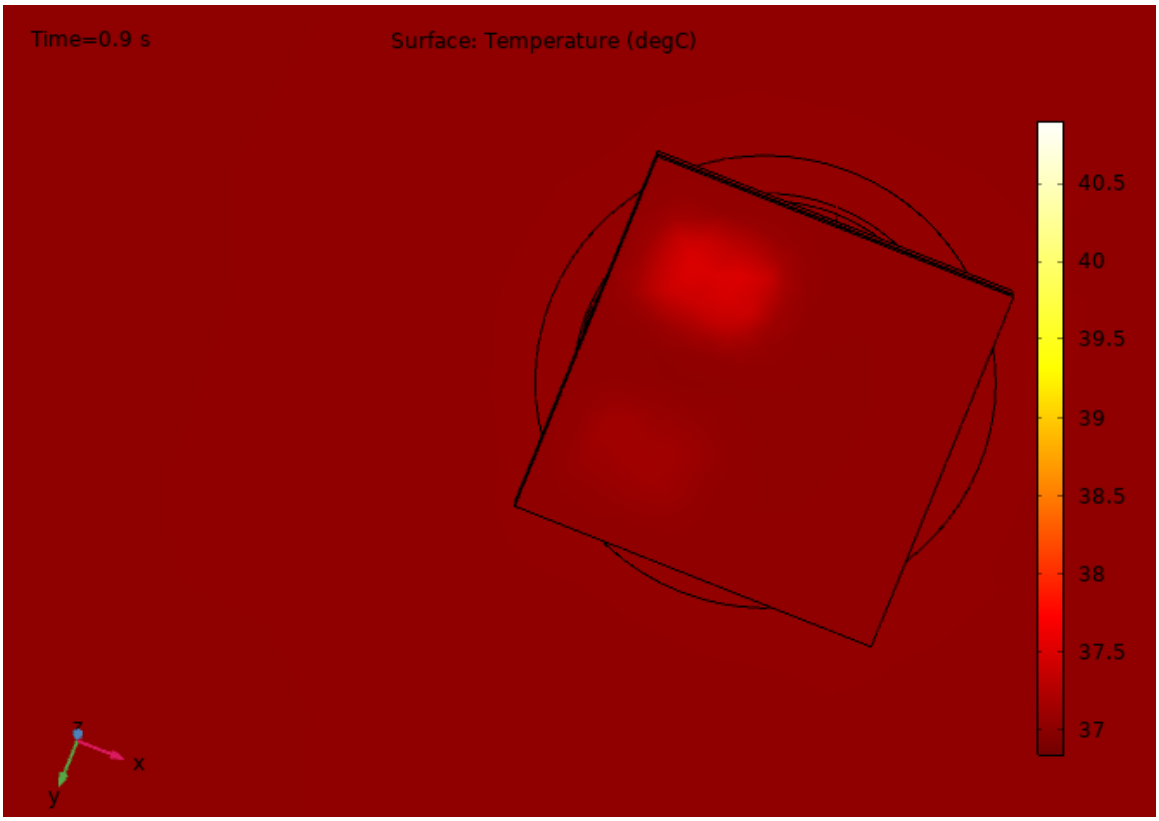


Figure 3.5: Conduction between two heat sources and the board

are set to help better present the results. In our final model, parameters are reset to emulate the human environment.

3.2.4 Coherence of the thermal effects

Even though we have taken into all the thermal effects into account, there is another phenomenon in physics that impacts the thermal distribution on the board, namely, coherence. The concept of coherence is related to the randomness of light sources intrinsically since heat is also a kind of light. In the COMSOL model, the coherence is caused by the randomness of electromagnetic fields. As paper [24] and [25] said, coherence of the filed can be characterized by a two point, two-time correlation function

$$< E(r_1, t_1), E(r_2, t_2) > \tag{3.5}$$

To help the reader better understand this conception, we can imagine there are two pinholes in an opaque screen on points S2 and S3. Then two diffracted fields are produced and we can see their interference at point S1 located at a same distance from both pinholes. The coherence problem can be solved by using spatial filters [8].

3.3 Chapter Summary

Conduction, convection, radiation are three main mechanisms for heating [23]. In this chapter, we analyzed different thermal effects in the COMSOL simulation environment, which coincides with the result shown in [26].

With the built COMSOL model, to make it close to the real experimental platform, in the following studies, we set the current as 20 A and the power of heat sources in the range of [0, 1]. The frequency of the AC of coil 1 is 13.56 MHZ, which is obtained from our earlier work [8]. The velocity of the fluid is 0.0022 m/s, which is to emulate the average velocity of blood. The induction thermal has an equation as

$$P = 0.1 \times 0.65 \times UI \tag{3.6}$$

where U is the voltage and I is the current of coil 1. This equation means, with 1% power transferred from coil 1 to coil 2, 65% of the power are turned into the thermal energy. With these parameters, we obtained the inputs and outputs of the system, which are used in Chapter 5 for system identification.

CHAPTER 4
SYSTEM IDENTIFICATION METHODS

In this chapter, the principle and structure of LSTM are introduced firstly since our plan is to use it for thermal dynamics modeling in the following two chapters. For comparison, the principle of RPBSID algorithm is also briefly introduced, which has been developed in our lab for capturing the thermal dynamics for IMDs [8]. The computing environment is also provided for reference.

4.1 LSTM in system identification

The LSTM was first introduced by [27] and gradually evolved to the widely used version. A general LSTM cell with memory function is composed of three gated units: forget gate, update gate, and output gate. It successfully solved the vanishing gradient problem existed before.

4.1.1 LSTM theory and Architecture

Figure 4.1 shows a single cell of LSTM, where x_k is the input of the network; h_k is the hidden layer state; σ and $tanh$ are respectively the sigmoid function and hyperbolic tangent function; W_f, W_e, W_i, W_c, W_o denotes the trainable weights for each layer and b_f, b_c, b_i,

b_o are the corresponding bias. The LSTM can be described by the following equations:

$$\begin{aligned} f_k &= \sigma(W_f \cdot [h_{k-1}, x_k] + b_f) \\ i_k &= \sigma(W_i \cdot h_{k-1} + b_i) \\ m_k &= tanh(W_c \cdot [h_{k-1} + b_c]) \\ g_k &= i_k \cdot m_k \\ c_k &= f_k * c_{k-1} + g_k \\ o_k &= \sigma(W_o \cdot [h_{k-1}, x_k] + b_o) \\ h_k &= o_k * tanh(c_k) \end{aligned} \tag{4.1}$$

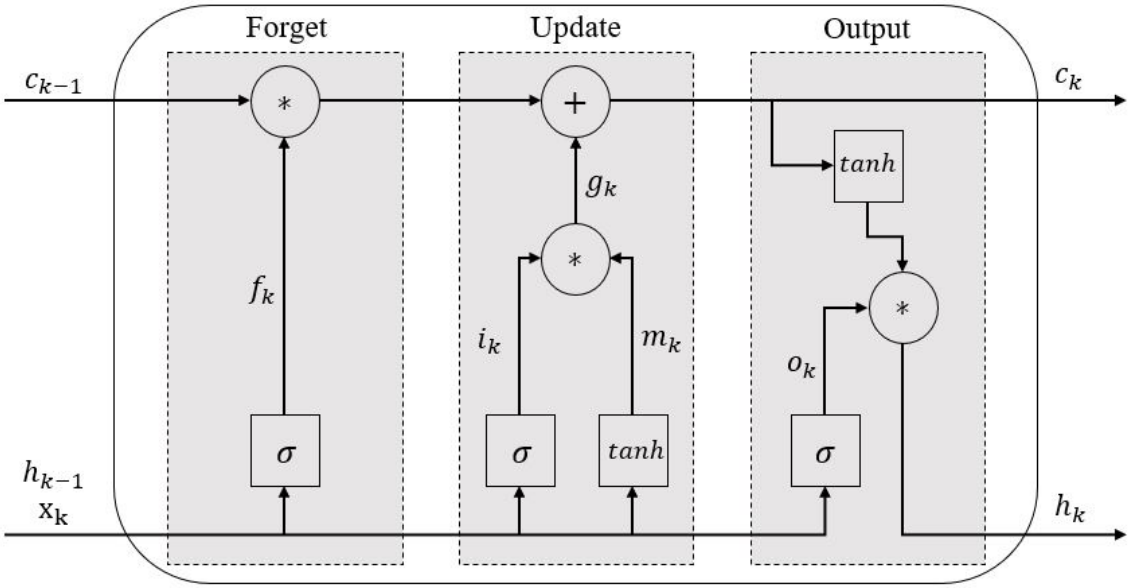


Figure 4.1: A single LSTM cell.

Different from one step ahead vector auto-regressive with exogenous inputs (VARX) models, which regards the output of the next time step as the linear combination of the past inputs and outputs, i.e.,

$$\hat{y}_{k|k-1} = \sum_{i=0}^{w} \alpha_i u_{k-i} + \sum_{i=1}^{w} y_{k-i}, \tag{4.2}$$

the mathematical representation of system identification for thermal dynamics using LSTM can be written as follows:

$$T(k) = f(T(k-1), T(k-2), ..., T(k-w); \\ u(k-1), u(k-2), ..., u(k-w)) \tag{4.3}$$

where $T(k)$ is the temperature at time step k; w is the window size; $u(k)$ is the input signal at time step k and $f(\cdot)$ is the nonlinear function that will be learned by the LSTM we designed. That means LSTM can learn the nonlinear function directly based on the universal-approximation theorem [28]. As a compromise between accuracy and time complexity, we designed an LSTM network consisting of two layers, one of which is LSTM layer and the other layer is a dense layer with six neurons as shown in Figure 4.2. This leads to a network with a dropout rate of 0.2 and a number of neurons of 500 that will be used in the following chapters. The Adam optimizer with mean squared error (MSE) optimization function is used to train the network.

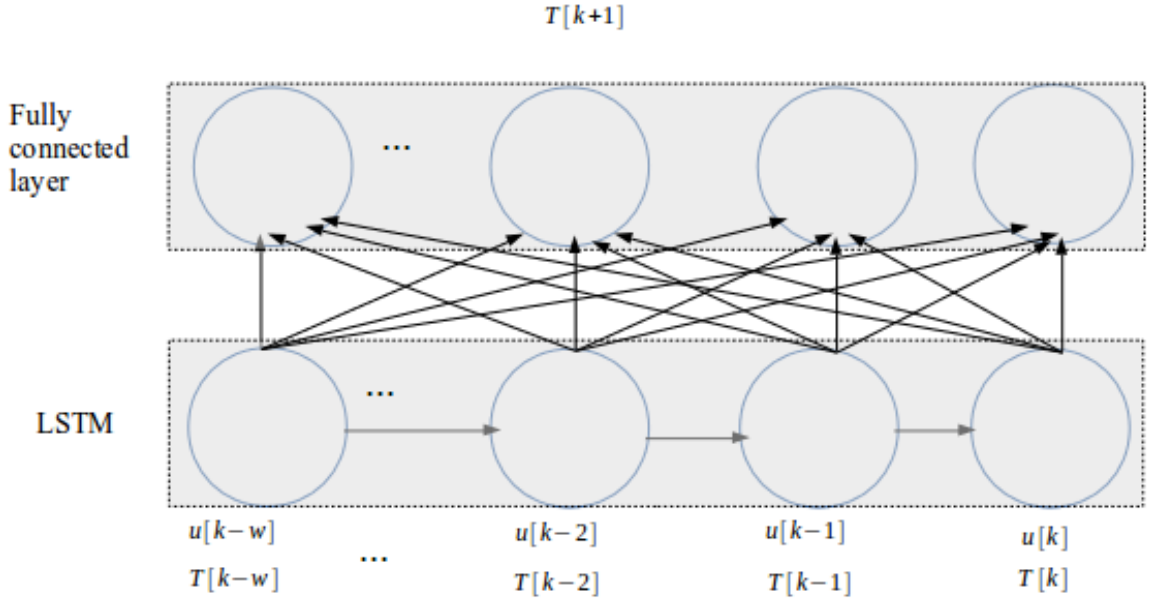


Figure 4.2: The overall structure of the LSTM network.

4.1.2 Online LSTM

For the online LSTM, to have a simple and efficient network, a one layer LSTM structure with 50 nodes is designed. The number of 50 is chosen because it provides the best results in terms of time complexity and accuracy on the test dataset (data obtained from our model in Chapter 3 with two different uniformly distributed random inputs). The node number we compared are 5, 50, 100 and 200. As shown in Table 4.1, when the number of nodes is further increased after 50, the accuracy does not improve prominently.

Table 4.1: PERFORMANCE ON TEST DATASET ($\times 10^{-3}$ °C)

# Nodes	**Mean Squared Error**
5	0.1565
50	0.0606
100	0.0536
200	0.0502

4.2 Recursive Predictor-Based Subspace Identification (RPBSID)

Suppose we have a system $\mathcal{S}$ with dynamics described by the following state-space model:

$$\mathcal{S}: \begin{cases} x_{k+1} = Ax_k + Bu_k + \omega_k \\ y_k = Cx_k + Du_k + v_k \end{cases} \tag{4.4}$$

where $x_k \in R^n$, $u_k \in R^r$, $y_k \in R^l$ are the state, input, output; $\omega_k \in R^l$ and $v_k \in R^l$ are respectively process noise and measurement noise and (A, B, C, D) are state-space matrices. A recursive system identification method is intended to solve the real-time identification problem. The following definition can help reader better understand this conception:

Recursive System Identification [29]: Given the system matrices $A_{k-1}, B_{k-1}, C_{k-1}, D_{k-1}$ and given the new input sample u_k and output y_k, find system matrices A_k, B_k, C_k, D_k in a similar state basis.

In essence, the RPBSID algorithm is a recursive system identification method. The principle of RPBSID algorithm is introduced in [29]. Distinguished from the conventional singular value decomposition (SVD)-based subspace identification method, three recursive least square (RLS) filters are used in this algorithm. The details of this algorithm is shown in **Algorithm** 1, where p is the size of the past window; f is the size of the future window; $\lambda_{1,2,3}$ and $\rho_{1,2,3}$ are updating parameters; n is the system order; I is the identity matrix; u_k and y_k are the input and output of this system respectively; S is a stationary selection matrix and P_k is the co-variance matrix. To save space, the explicit definition of the remaining matrices in this algorithm can be found in Appendix A.

As an existing work in our lab, the RPBSID algorithm in [8] is used for modeling the system dynamics with our data. Firstly, spatial filtering methods [30] are utilized to divide the data into two components based on spatial dependency. After filtering out the spatial component of the data, RPBSID methods are applied to the non-spatial components by using a vector auto regressive with exogenous inputs (VARX) model [31]. Then, a kernel recursive least squares (KRLS) filter [32] with surprise criterion (SC) is applied to the spatial components to obtain the spatial predictor. The combination of the spatial and non-spatial predictions are the ultimate prediction of the system.

4.3 Computing Environment

A personal computer with Ubuntu 16.04.4, python version 3.5.2, tensorflow 1.4.0 and CUDA version 8.0.6 is used for the training and testing process, which is sufficient due to the simple structure of the network. The same computer and hardware environment will be used for the comparison of different algorithms.

4.4 Chapter Summary

This chapter has presented the principle of two algorithms, (i.e., LSTM and RPBSID) that are used in our research. The structure of the LSTM was determined by comparing the

Algorithm 1 The RPBSID algorithm

input p, f, n, S, Θ_{-1}, $\lambda_{1,2,3}$, $\rho_{1,2,3}$
require $n > 0, n/l \leq f \leq p, 0 \ll \lambda_{1,2,3} \leq 1, \rho_{1,2,3} > 0$
init $P_{-1} = (1/\rho_1)I$, $M_{-1} = (1/\rho_2 I)$, $N_{-1} = (1/\rho_3 I)$, $\tilde{E}_{-1}$ from Θ_{-1}
for k = 0, 1,2,...

input u_k, y_k
Step 1: Update the Markov parameters
Update RLS:

$$P_k = \frac{1}{\lambda_1} P_{k-1} - \frac{1}{\lambda_1} P_{k-1}\phi_k(\lambda_1 I + \phi_k^T P_{k-1}\phi_k)^{-1}\phi_k^T P_{k-1} \quad \textbf{[A.1]} \tag{4.5}$$

$$\tilde{E}_k = \tilde{E}_{k-1} + (y_k - \tilde{E}_{k-1}\phi_k)\phi_k^T P_k \quad \textbf{[A.3]} \tag{4.6}$$

Step 2: Estimate the state vector
Construct the matrices $\tilde{\Gamma}_k$ **[A.5]** and $\tilde{\Upsilon}_k$ **[A.6]**
Construct Hanker matrix $\tilde{H}_k$ **[A.7]**
Compute: $\hat{x}_k = S\tilde{H}_k^{-1}(\tilde{\Gamma}_k \bar{u}_{k-p,p} + \tilde{\Upsilon}_k \bar{y}_{k-p,p})$ **[A.2]**
Step 3: Update the system matrices:
Update RLS:

$$M_k = \frac{1}{\lambda_2} M_{k-1} - \frac{1}{\lambda_2} M_{k-1}\psi_{k-1}(\lambda_2 I + \psi_{k-1}^T M_{k-1}\psi_{k-1})^{-1}\psi_{k-1}^T M_{k-1} \quad \textbf{[A.11]} \tag{4.7}$$

$$\Theta_k^{(y)} = \Theta_{k-1}^{(y)} + (y_{k-1} - \Theta_{k-1}^{(y)}\psi_{k-1})\psi_{k-1}^T M_k \quad \textbf{[A.10]} \tag{4.8}$$

Compute $e_{k-1} = y_{k-1} - \Theta_k^{(y)}\psi_{k-1}$
Update RLS:

$$N_k = \frac{1}{\lambda_3} N_{k-1} - \frac{1}{\lambda_3} N_{k-1}\varphi_{k-1}(\lambda_3 I + \varphi_{k-1}^T N_{k-1}\varphi_{k-1})^{-1}\varphi_{k-1}^T N_{k-1} \quad \textbf{[A.12]} \tag{4.9}$$

$$\Theta_k^{(x)} = \Theta_{k-1}^{(x)} + (x_k - \Theta_{k-1}^{(x)}\varphi_{k-1})\varphi_{k-1}^T N_k \quad \textbf{[A.9]} \tag{4.10}$$

performance under different hyper-parameters and compromising between accuracy and efficiency. These algorithms will be applied to the IMD thermal dynamics data in the following two chapters.

CHAPTER 5
EVALUATION OF THERMAL DYNAMIC MODELS USING SIMULATION STUDIES

To evaluate the effectiveness and robustness of the proposed thermal dynamic models, several experiments are carried out using both simulation and *in vitro* experiments. This chapter presents the results based on the simulation studies.

5.1 Data Collection

The COMSOL model presented in Chapter 3 is used to provide simulation data. Eight scenarios, as listed in Table 5.1 are considered in the simulation study to test the effectiveness of the algorithms. To test the robustness, extreme on-off inputs such as "OD1H", "ODS1H", "OD1M" and etc. are provided to the two heat sources on TMTV, in which one or both inputs have sudden change during simulation. In all on-off input scenarios, inputs have the same 10% duty cycle unless specifically stated. The plot of the power inputs are separately shown in Figure 5.1 and Figure 5.2.

The output temperatures at S1-S6 are shown in Figure 5.3 and Figure 5.4 for two simulation scenarios. As shown in both figures, S2 and S3 have higher temperatures than other positions because they are at the heat sources. The temperature at S1 is influenced by the coherence of thermal radiation, thus it has a more curvy trajectory compared to temperatures at S4-S6. For other simulation scenarios, similar trends are observed.

5.2 Data Pre-processing

As previously mentioned in Chapter 4, the RPBSID algorithm utilizes a spatial filter to separate the original data into a 1×1 spatial vector and a 6×1 non-spatial vectors for

Table 5.1: EXPLANATION OF COMSOL STUDIES

Scenarios	**Explanation**
RandDiff2	Two inputs (H1, H2) are different uniformly distributed random signal from range [0,1] (different from RandDiff)
RandSame	Two inputs (H1, H2) are same with a uniformly distributed random signal from range [0,1]
OD1H	Increase the input amplitude of H1 from 0.5W to 0.75W while decreasing the input amplitude of H2 from 0.75W to 0.5W
OD1S	Increase the input amplitude of H1 from 0.5W to 0.75W while decreasing the input amplitude of H2 from 0.75W to 0.5W (sudden change)
OSD1H	Increase the input amplitude of H1 from 0.5W to 0.75W while decreasing the input amplitude of H2 from 0.5W to 0.25W
ODS1H	Increase the input amplitude of H1 from 0.5W to 0.75W while increasing the input amplitude of H2 from 0.25W to 0.75W
OD1M (dc=10%)	Decease the input amplitude of H1 from 0.75W to 0.375W while increasing the input amplitude of H2 from 0.375W to 1.5W with duty cycle 10
OD1M (dc=50%)	Decease the input amplitude of H1 from 0.75W to 0.375W while increasing the input amplitude of H2 from 0.375W to 1.5W with duty cycle 50

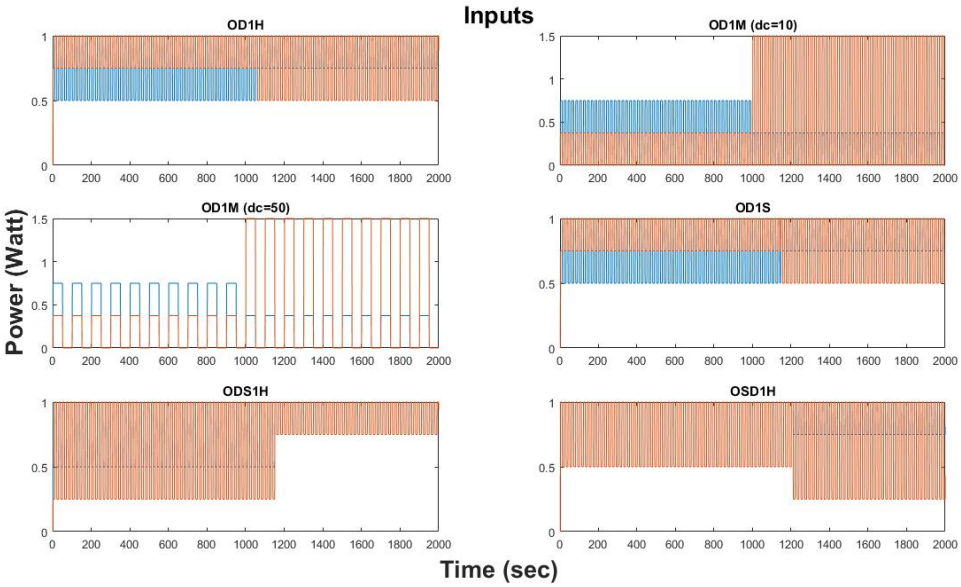


Figure 5.1: Plots of power inputs for each on-off scenario

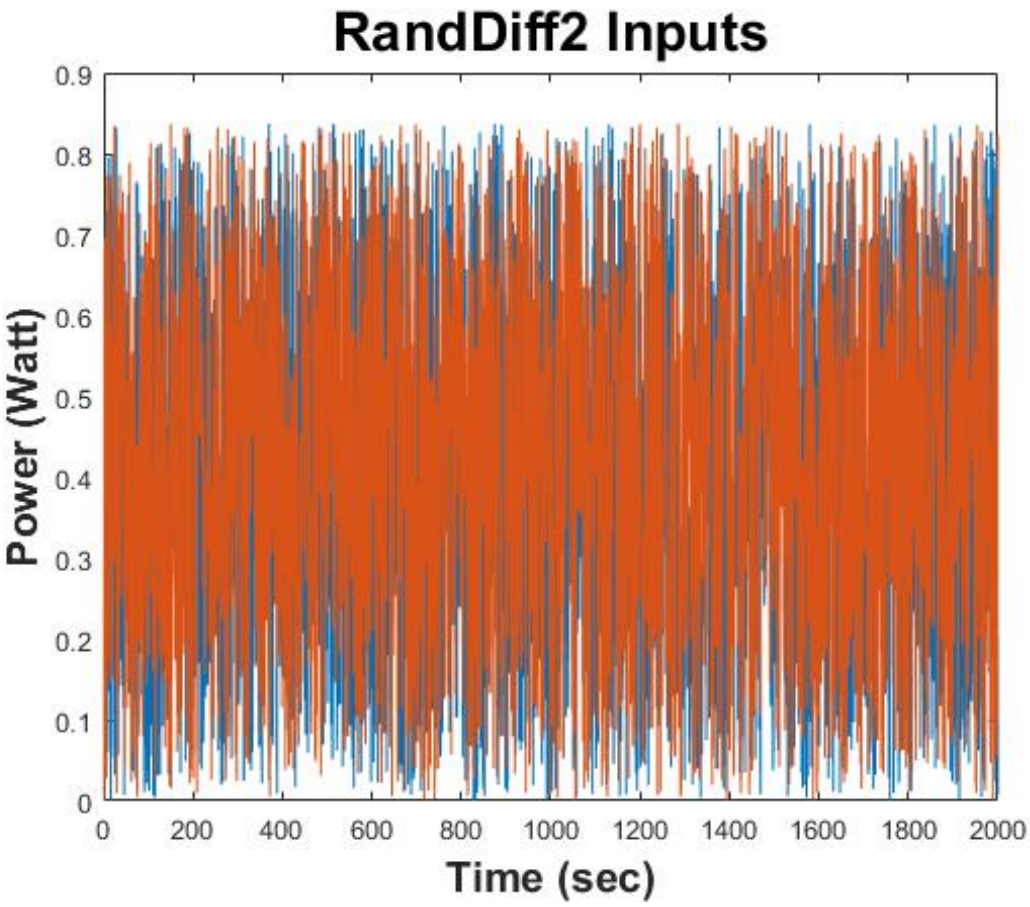


Figure 5.2: Plot of RandDiff2 inputs

the simulation data. The same method is used here to pre-process the simulation data. The principle of spatial filter is introduced in [30]. In order to train the neural network and obtain a suitable model, a training dataset with two uniformly distributed random inputs, namely, ”RandDiff”, are used and this ”RandDiff” dataset is not the same as ”RandDiff2” in Table 5.1. Batch normalization is used to improve the performance. As a result, the standard deviation and the mean are continuously updated on each time horizon of inputs, which is believed to smooth the objective function and thus improve the performance [33]. These data pre-processing methods will also be used for experiment data in Chapter 6.

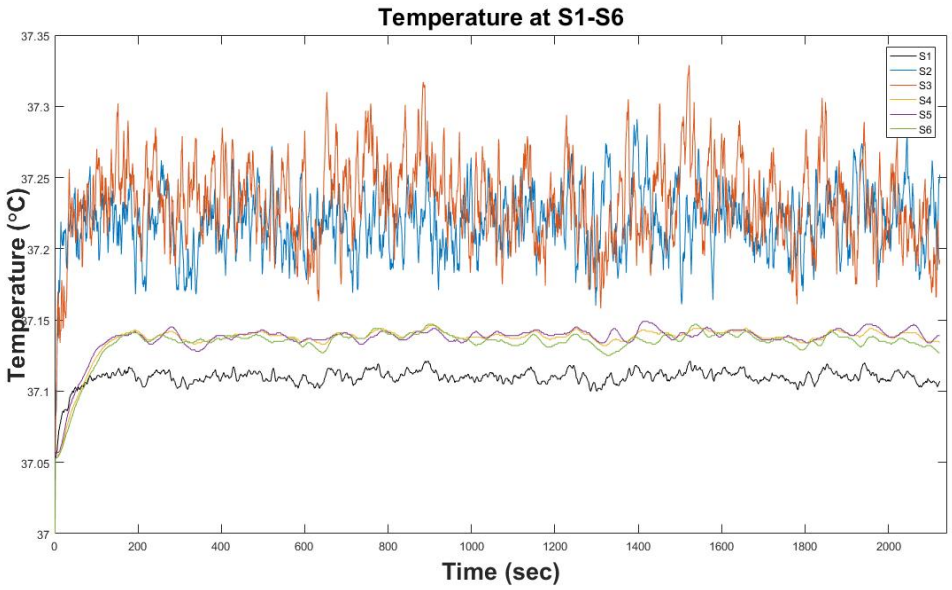


Figure 5.3: Temperature trend chart: ”RandDiff2”

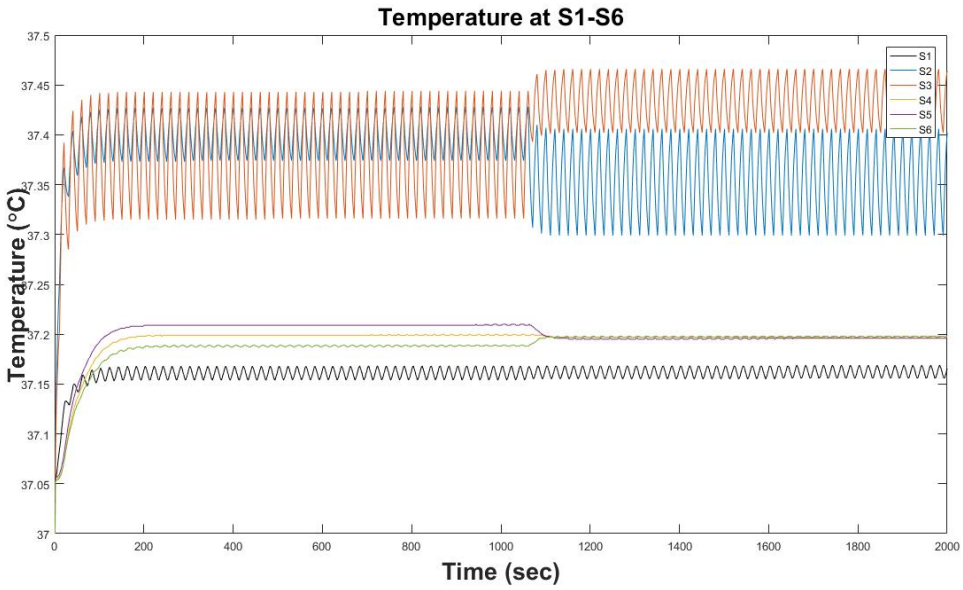


Figure 5.4: Temperature trend chart: ”OD1H”

5.3 Evaluation Metrics

Mean Squared Error (MSE) is a widely used metric when weighing the performance of a machine learning algorithm and can be formulated as

$$MSE = \frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{y}_t)^2 \tag{5.1}$$

where $\hat{y}_t$ is the prediction result in time t.

Nonetheless, if the change of the output is negligible, using only MSE value as a performance metric is not sufficient. So best fit rate (BFR) is also utilized as an auxiliary measurement, the definition of which is

$$BFR = max\left\{1 - \sqrt{\frac{\sum_{t=1}^{N}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{N}(y_t - \bar{y}_t)^2}}, 0\right\} \tag{5.2}$$

where $\bar{y}$ defines the mean of the real output and $\hat{y}$ is the predicted output. As its definition suggests, the higher the BFR value, the better the prediction performance.

To compare the prediction performance of different methods, both MSE and BFR are used in this chapter and Chapter 6.

5.4 Results and Discussion

For the LSTM algorithm, the process is

- Step 1: train the model using RandDiff dataset because it can capture all of the frequency of the system.
- Step 2: use the model from step 1 to predict online data with inputs stated in Table 5.1.

In order to have an intuitive understanding, we presented the parameters of three algorithms in Table 5.2, in which ”/” means no parameters.

Table 5.2: PARAMETERS OF THREE ALGORITHMS

Scenarios	**LSTM**	**Online LSTM**	**RPBSID**
number of layer	1	1	/
number of hidden nodes	500	50	/
optimizer configure	adam	adam	/
number of input	8	8	8
number of output	6	6	6
window size	10	10	10
Activation	linear	linear	/
Forgetting factor	/	/	$\lambda_1, \lambda_2, \lambda_3$

Considering that RPBSID has an adaptive phase before it can really capture the dynamics of the system, we start to calculate the MSE and BFR of prediction results of the RPBSID algorithm after 400 data samples for all the following experiments. Since there are six outputs, the average MSE and average BFR of six outputs are calculated for each algorithm and are summarized in Table 5.3 and Table 5.4 respectively, where "RPBSID*" means the RPBSID algorithm after convergence is achieved.

Table 5.3: COMPARISON OF DIFFERENT ALGORITHMS BASED ON THE SIMULATION DATA: MSE ($\times 10^{-3}$ °C)

Algorithms	**LSTM**	**Online-LSTM**	**RPBSID***
RandDiff	0.1585	0.0552	0.0341
RandSame	0.1668	0.0443	0.0274
OD1H	0.5687	0.0383	0.0572
OD1S	0.5738	0.0458	0.0386
OSD1H	0.9297	0.0565	0.3036
ODS1H	0.8293	0.0514	0.0462
OD1M (dc=10%)	2.6222	0.1538	4.2680
OD1M (dc=50%)	1.2760	0.1041	0.1886

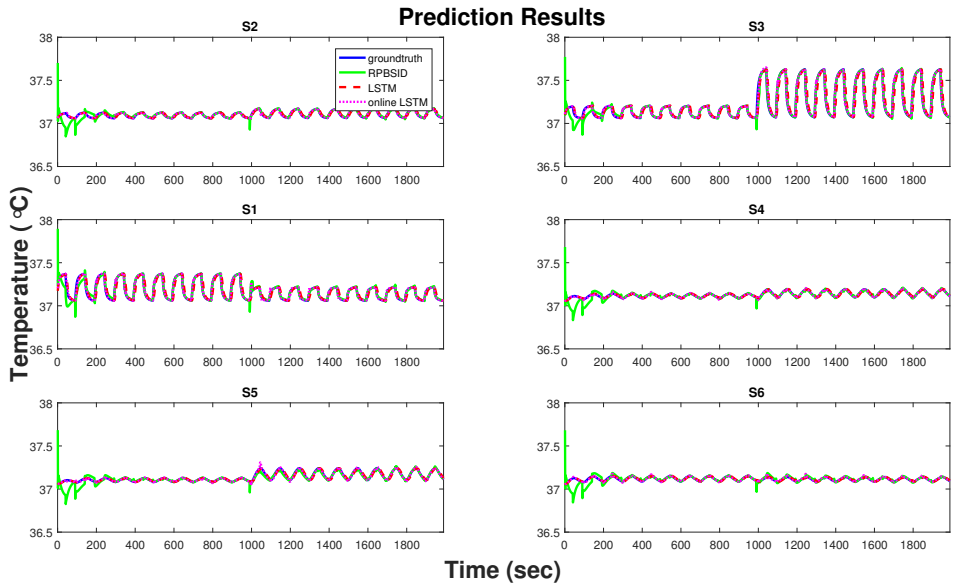
As shown in Table 5.3 and Table 5.4, the online LSTM has a significantly better performance on BFR compared to offline LSTM and RPBSID. From the aspect of MSE, RPBSID

Table 5.4: COMPARISON OF DIFFERENT ALGORITHMS BASED ON THE SIMULATION DATA: BFR (%)

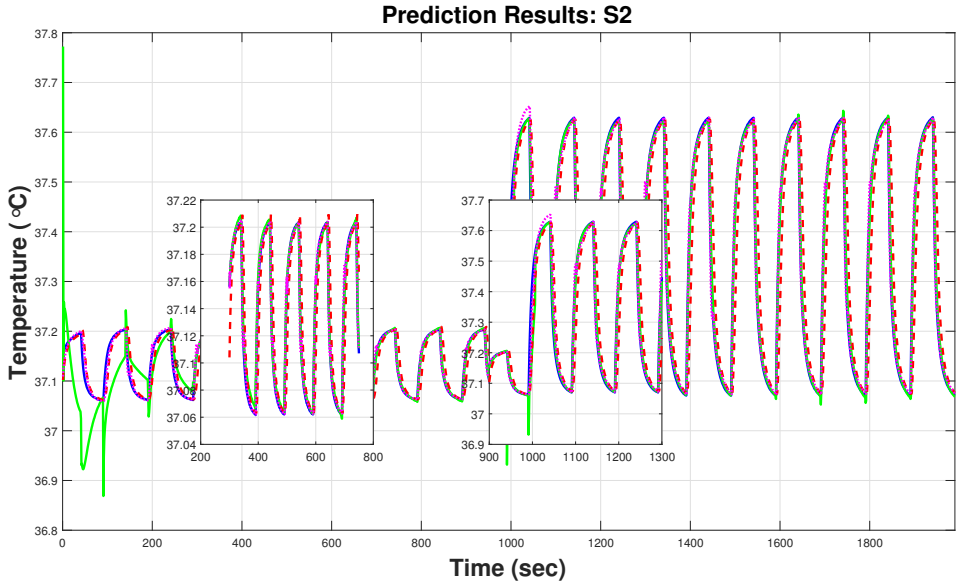
Algorithms	**LSTM**	**Online-LSTM**	**RPBSID***
RandDiff	73.130	73.666	23.533
RandSame	72.148	83.541	35.112
OD1H	59.632	61.721	32.796
OD1S	59.931	56.960	45.919
OSD1H	47.937	72.472	13.357
ODS1H	69.153	78.550	72.045
OD1M (dc=10%)	30.223	67.423	8.5380
OD1M (dc=50%)	74.641	87.885	76.153

outperforms offline LSTM after convergence, which is reasonable since RPBSID is an online system identification method. On the other hand, the duty cycle of the input signal also has an influence on the prediction performance. "OD1M" with 50% duty cycle has a better performance than "OD1M" with a duty cycle of 10% considering both MSE and BFR. This influence is shown in all three algorithms, which is a result of the difficulty in capturing the dynamics of a fast time-varying dynamical system for both LSTM-based and conventional methods.

The prediction results of scenario "OD1M (dc=50)" are used as an example for further illustration, as shown in Figure 5.5. Particularly, the prediction results at S2 are displayed in Figure 5.5 (b) to show the difference of the three algorithms. As shown in Figure 5.5 (b), the RPBSID algorithm has a poor performance for the first 200 seconds while the LSTM-based algorithm can capture the dynamics from the start. From 300 second to 750 second, all three algorithms perform quite well. At 1000 second, there is an abrupt magnitude lifting of the temperature because of the increase of power input at the corresponding heat source H2, which, however, doesn't impact the tracking ability of the LSTM. At the same time, for the RPBSID algorithm, even though there is a sharp decrease in temperature, it adapts to track the real trajectory again quickly.



(a)



(b)

Figure 5.5: (a) Simulation results for "OD1M (dc=50)", (b) Simulation result for S2 (zoomed-in)

The LSTM algorithm is not indefectible in that there is a slight offset of the prediction results for the offline LSTM algorithm compared to the remainders. In addition, online LSTM faces a challenge because of its time complexity. The time complexity of the RPB-SID algorithm for each sample is about 0.0001 seconds while the online LSTM algorithm

has a time complexity of 0.01 seconds.

5.5 Chapter Summary

In this chapter, we compared the performance of the RPBSID, the offline LSTM and the online LSTM algorithms and concluded each algorithm's strengths and weaknesses. Different from the regular machine learning process (i.e., train and test for the data with similar distribution), we trained the model with data obtained from our model in Chapter 3 with two different uniformly distributed random inputs (i.e., "RandDiff") and tested on other scenarios' data. We presented the superiority of the LSTM-based system identification methods when there is an abrupt change of the input signals. As demonstrated in this chapter, the LSTM algorithm has a good ability of learning time dependent systems due to its memory modules. One reason why LSTM is not widely used in control field is its time complexity. However, a higher performance GPU can be used to reduce this limitation.

CHAPTER 6
EVALUATION OF THERMAL DYNAMIC MODELS USING *IN VITRO* EXPERIMENT

In this chapter, the proposed thermal dynamic model is evaluated using *in vitro* experiments that use a temperature monitoring and management test vehicle developed in our lab.

6.1 Experimental Setup and Data Collection

The overall experimental setup [8] is displayed in Figure 6.1. Since the wireless power transfer part is still under construction, the TMTV is powered directly by a power supply in this experiment. The TMTV board is submerged into a water tank which aims to model the blood circulation. Then a nRF52 board is connected to a personal computer with a created Matlab GUI to receive the signal transmitted from the TMTV board. On the TMTV board, there are two heat sources that receive power input from power supply and six temperature sensors (LMT70) to monitor the temperature variations on the board. To emulate the blood flow, a marine pump is placed into the water-filled container to create water circulation and the sponge is used to ensure a uniform water flow (i.e., blood flow). A close-up observation of the board is shown in Figure 6.2.

Four groups of studies have been deployed to test the performance of three algorithms. A description of the four studies is presented in Table 6.1. Heat sources on TMTV admit pulse width modulation (PWM) inputs within the range of $[0, 10000]$ where 0 means a PWM signal with 100% duty cycle and 10000 means a PWM signal with 0% duty cycle. PWM inputs in all experiments have the same duty cycle of 10% [8]. The temperature distributions at S1-S6 of scenario "RandDiff" and "OD1H" are displayed in Figure 6.3 and Figure 6.4 respectively.

Similar to Chapter 5, the "RandDiff" dataset is used as the training set and the remaining

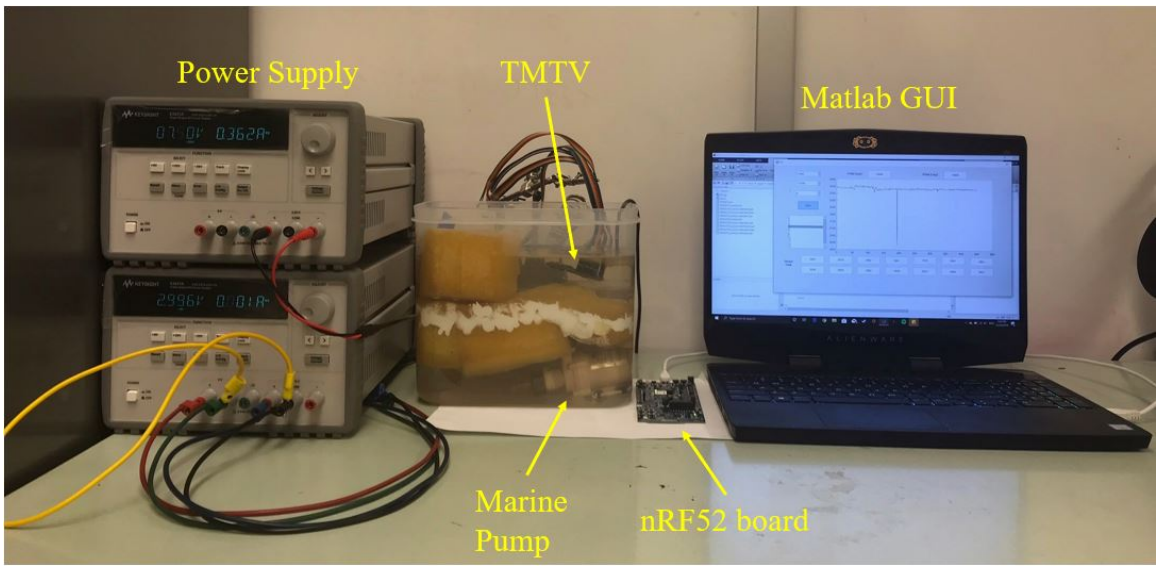


Figure 6.1: Experimental setup

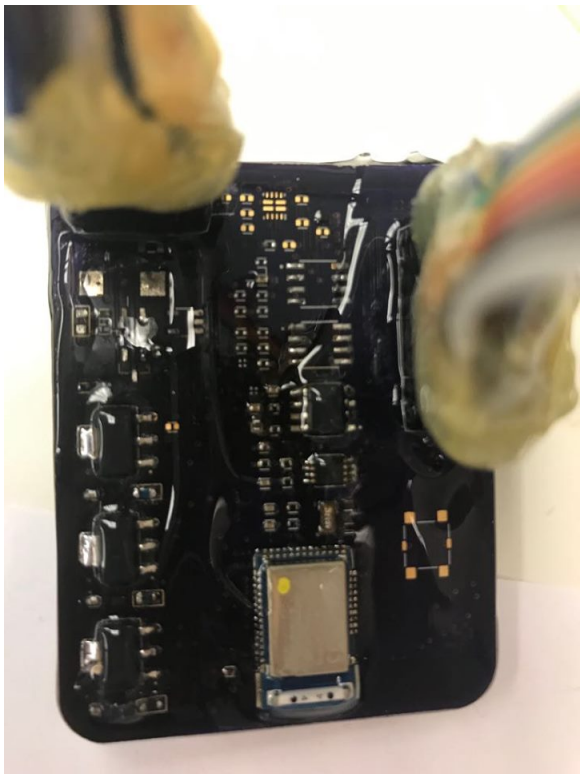


Figure 6.2: TMTV board

four scenarios are used as test dataset. For the RPBSID algorithm, the data obtained is separated into non-spatial part and spatial part; specially, a fixed-lag Kalman filter with zero-mean Gaussian process and measurement noise is used for the experiment data. For

Table 6.1: EXPLANATION OF EXPERIMENT STUDIES

Scenarios	**Explanation**
OD1H	Increase the PWM input of H1 from 5000 to 7500 while decreasing the PWM input of H2 from 7500 to 5000
OD1S	Increase the PWM input of H1 from 5000 to 7500 while decreasing the PWM input of H2 from 7500 to 5000 (sudden change)
ODS1H	Change the PWM inputs of H1 from 5000 to 7500 while changing the PWM input of H2 from 2500 to 7500
OSD1H	Increase the PWM input of H1 from 5000 to 7500 while decreasing the PWM input of H2 from 5000 to 2500

all the experiments, a training data is used for batch pre-processing such that the updating can be initiated after 10 seconds when the input and output are obtained.

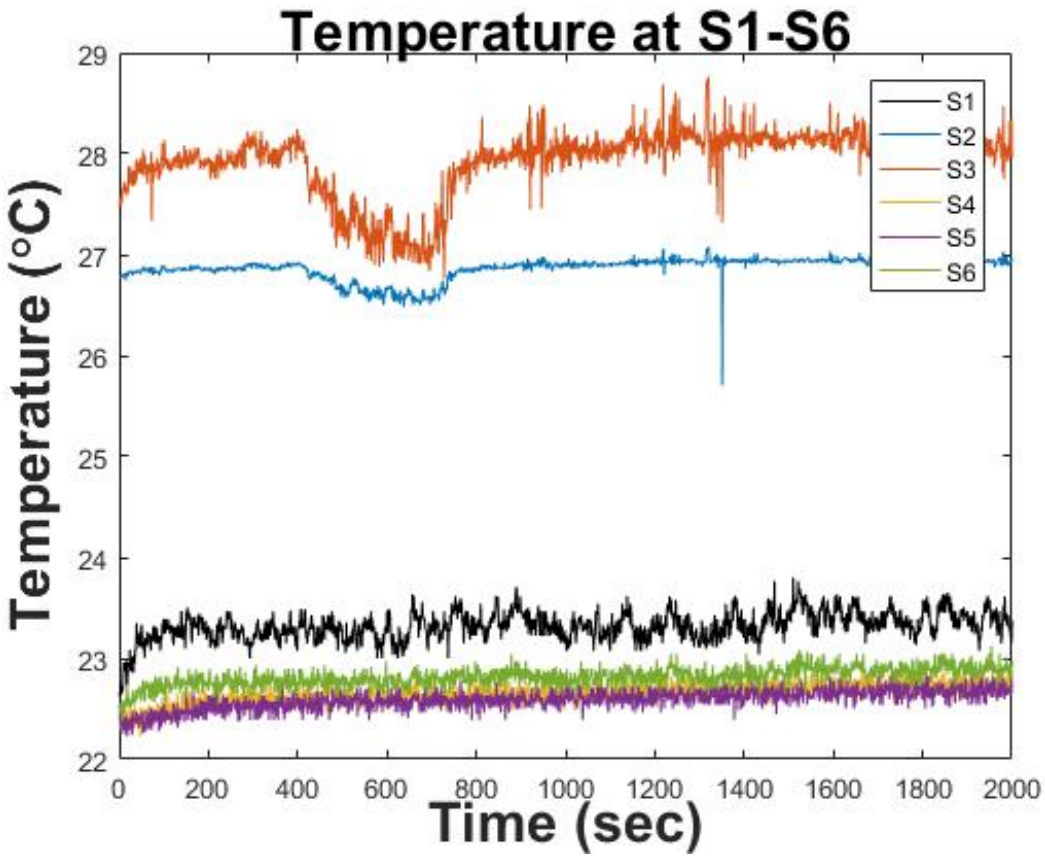


Figure 6.3: Temperature trend chart: ”RandDiff”

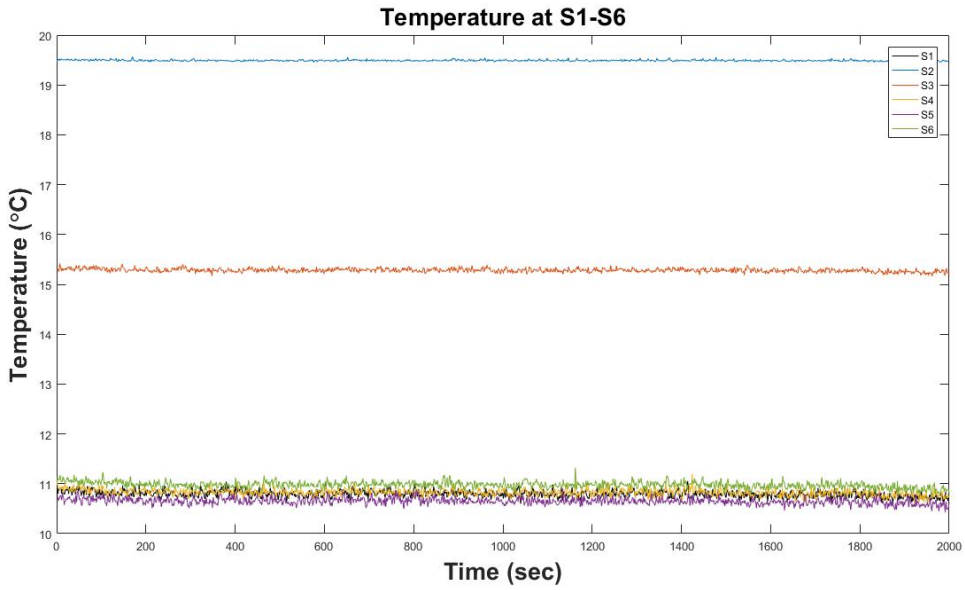


Figure 6.4: Temperature trend chart: ”OD1H”

6.2 Results and Analysis

Table 6.2 and Table 6.3 display the performance of three algorithms. Considering both the average MSE and the average BFR values of six outputs, we conduce that the online-LSTM algorithm has a mildly-improved performance compared to offline LSTM algorithm and offline LSTM algorithm outperform RPBSID algorithm for experiment data in scenarios ”OD1H”, ”ODS1H” and ”OSD1H”. In Table 6.3, the BFR values of scenarios ”OD1H” and ”OSD1H” for the RPBSID algorithm are 0, which indicates the invalidity of the model [34].

Table 6.2: COMPARISON OF DIFFERENT ALGORITHMS BASED ON HE EXPERIMENTAL DATA: MSE (°C)

Algorithms	**LSTM**	**Online-LSTM**	**RPBSID***
OD1H	0.0033	0.0026	0.3780
OD1S	0.0441	0.0143	0.0100
ODS1H	0.0370	0.0104	0.0080
OSD1H	0.0041	0.0030	0.0650

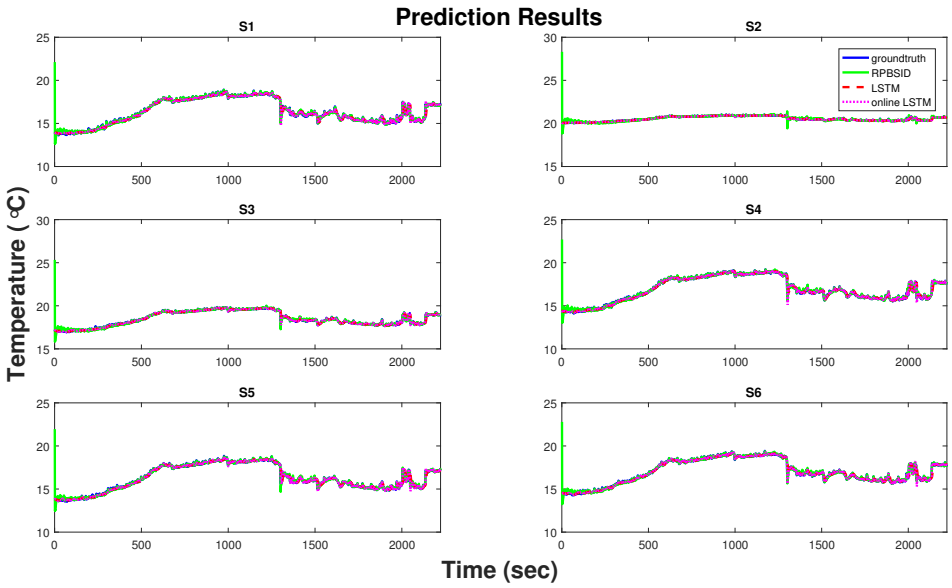
Table 6.3: COMPARISON OF DIFFERENT ALGORITHMS BASED ON HE EXPERIMENTAL DATA: BFR (%)

Algorithms	**LSTM**	**Online-LSTM**	**RPBSID***
OD1H	53.29	56.41	0.000
OD1S	89.33	91.23	82.11
ODS1H	94.36	95.79	87.73
OSD1H	63.07	64.70	0.000

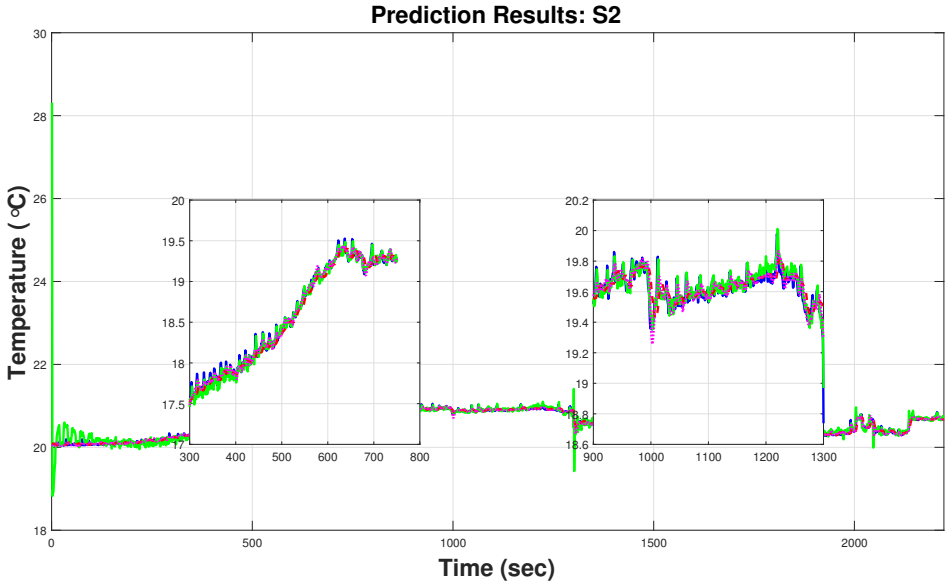
Figure 6.5 shows the prediction results of all algorithms for the scenario "ODS1H", from which, we can see that all three algorithms can capture the behavior of the system in general. The RPBSID algorithm and the online LSTM algorithm are better at tracking the dynamics adaptively while the offline LSTM algorithm has a distinct delay when tracking the time-varying dynamics of the system.

6.3 Chapter Summary

Examining both the simulation and experiment results, an interesting phenomenon can be observed, that is, for the same data, the LSTM algorithm and the RPBSID algorithm have the same trend in learning the temperatures at S1-S6. That is to say, when the LSTM algorithm can't learn well at positions S2 and S3 but performs well at positions S1, S4, S5 and S6, the RPBSID algorithm shows the same behavior and vice versa. Our explanation is that the regression methods used in both algorithms are similar. The main difference in the regression methods used in the algorithms is that the RPBSID algorithm is based on state space. Thus, the RPBSID algorithm can only find a local minimum, since only a local area of information are available for each time instance. However, the LSTM algorithm can learn the dynamics of the whole system since the whole dataset with two different uniformly distributed random inputs (i.e., "RandDiff") is used and is more likely to find the global minimum of the objective function. Prediction results for experiment data is not as accurate as simulation data because of the noise from the environment and the experiment



(a)



(b)

Figure 6.5: (a) Experiment results for "ODS1H", (b) Experiment result for S2 (zoomed-in)

setup.

CHAPTER 7
CONCLUSIONS AND FUTURE WORK

7.1 Contributions

In this thesis, modelling of thermal dynamics of IMD system is explored. With consideration of all the thermal effects of the wireless charging system, we developed LSTM-based system identification algorithms and compared the performance of three different algorithms (i.e., LSTM, Online LSTM, and RPBSID) using both the simulation data and experiment data. Both simulation results and experimental results demonstrate the advantage of LSTM algorithm in prediction. According to the simulation and experiment results, LSTM algorithms have performance advantages compared to conventional matrix-based system identification methods. The advantages of LSTM are (1) filtering and complex preprocessing for data are not required for the LSTM algorithm; and (2) once the model is trained, all of the dynamic information hidden in this system is learned by the LSTM. However, the LSTM has one disadvantage compared to the RPBSID algorithm, specifically, in the adaptive predicting phase. The online LSTM can learn adaptively with a moving window but with a high time complexity. All three algorithms involve modulating parameters to get a better performance. It is probable that with the development of hardware and quantum computer, implementing LSTM in implantable medical devices will be realizable.

7.2 Future Works

In the future, the proposed thermal dynamic model will be further evaluated using a wireless TMTV so that the effect of wireless power transfer can be included. In addition, model predictive control will be developed based on the LSTM thermal dynamic models to achieve real-time thermal dynamics control of IMDs.

Appendices

APPENDIX A

RPBSID ALGORITHM SYMBOLS

The symbols in the RPBSID algorithm (I is the identity matrix as commonly used):

$$\phi_k = \begin{bmatrix} \bar{u}_{k-p,p}^T & u_k^T & \bar{y}_{k-p,p}^T \end{bmatrix}^T \tag{A.1}$$

where

$$\bar{u}_{k-p,p} = \begin{bmatrix} u_{k-p} & u_{k-p+1} & \cdots & u_{k-1} \end{bmatrix}^T,\ \bar{y}_{k-p,p} = \begin{bmatrix} y_{k-p} & y_{k-p+1} & \cdots & y_{k-1} \end{bmatrix}^T. \tag{A.2}$$

$$\tilde{E} \triangleq \begin{bmatrix} \tilde{E}^{u_{k-p}} & \cdots & \tilde{E}^{u_k} & \tilde{E}^{y_{k-p}} & \cdots & \tilde{E}^{y_k} \end{bmatrix} \tag{A.3}$$

where

$$\tilde{E}^{(u_{k-i})} = \begin{cases} D, \text{ if } i = 0 \\ C\tilde{A}^{i-1}\tilde{B}, \text{ if } i > 0 \end{cases},\ \tilde{E}^{(y_{k-i})} = C\tilde{A}^{i-1}K \tag{A.4}$$

and $\tilde{A} = A - KC,\ \tilde{B} = B - KD$ and K is the Kalman gain matrix.

$$\tilde{\Gamma} = \begin{bmatrix} \tilde{E}^{u_{k-p}} & \tilde{E}^{u_{k-p+1}} & \cdots & \tilde{E}^{u_{k-p+f-1}} & \cdots & \tilde{E}^{u_{k-1}} \\ 0 & \tilde{E}^{u_{k-p}} & \cdots & \tilde{E}^{u_{k-p+f-2}} & \cdots & \tilde{E}^{u_{k-2}} \\ \vdots & \ddots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \tilde{E}^{u_{k-p}} & \cdots & \tilde{E}^{u_{k-f}} \end{bmatrix} \tag{A.5}$$

$$\Upsilon = \begin{bmatrix} \tilde{E}^{y_{k-p}} & \tilde{E}^{y_{k-p+1}} & \cdots & \tilde{E}^{y_{k-p+f-1}} & \cdots & \tilde{E}^{y_{k-1}} \\ 0 & \tilde{E}^{y_{k-p}} & \cdots & \tilde{E}^{y_{k-p+f-2}} & \cdots & \tilde{E}^{y_{k-2}} \\ \vdots & \ddots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \tilde{E}^{y_{k-p}} & \cdots & \tilde{E}^{y_{k-f}} \end{bmatrix} \tag{A.6}$$

$$\tilde{H} = \begin{bmatrix} I & 0 & \cdots & 0 \\ -CK & I & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ -C\tilde{A}^{f-2}K & -C\tilde{A}^{f-3}K & \cdots & I \end{bmatrix} \tag{A.7}$$

$$\psi_k = \begin{bmatrix} \hat{x}_k^T & u_k^T \end{bmatrix}^T \tag{A.8}$$

$$\Theta_k^{(x)} = [A_k, B_k, K_k], \tag{A.9}$$

$$\Theta_k^{(y)} = [C_k, D_k] \tag{A.10}$$

where A_k, B_k, C_k, D_k are the state space matrices and K_k is the Kalman gain matrix at step k.

$$e_k = y_k - \Theta_k^{(y)}\psi_k \tag{A.11}$$

$$\varphi_k = \begin{bmatrix} \hat{x}_k^T & u_k^T & \hat{e}_k^T \end{bmatrix}^T \tag{A.12}$$

REFERENCES

[1] Y.-H. Joung, "Development of implantable medical devices: From an engineering perspective," *International neurourology journal*, vol. 17, pp. 98–106, Sep. 2013.

[2] T. Fujii and Y. Ibata, "Effects of heating on electrical activities of guinea pig olfactory cortical slices," *Pflügers Archiv*, vol. 392, pp. 257–260, 1982.

[3] "Spreading depression elicited by thermal effects of ultrasonic irradiation of cerebral cortex in rats," *Journal of Neurobiology*, vol. 8, no. 4, pp. 381–393, 1977.

[4] G. Lazzi, "Thermal effects of bioimplants," *IEEE engineering in medicine and biology magazine : The quarterly magazine of the Engineering in Medicine Biology Society*, vol. 24, pp. 75–81, Jan. 2005.

[5] S. Kim, P. Tathireddy, R. Normann, and F. Solzbacher, "Thermal impact of an active 3-d microelectrode array implanted in the brain," *IEEE transactions on neural systems and rehabilitation engineering : A publication of the IEEE Engineering in Medicine and Biology Society*, vol. 15, pp. 493–501, Jan. 2008.

[6] R. Chai, Y. Zhang, and M. Ghovanloo, "Joint power and thermal management for implantable devices," in *2015 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 2015, pp. 1–4.

[7] R. Chai and Y. Zhang, "Adaptive thermal management of implantable device," *IEEE Sensors Journal*, vol. 19, no. 3, pp. 1176–1185, 2019.

[8] A. Ermis, Y. Lai, X. Pan, R. Chai, and Y. Zhang, "Recursive subspace identification for online thermal management of implantable devices," in *2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2019, pp. 944–949.

[9] M. Hardt, T. Ma, and B. Recht, "Gradient descent learns linear dynamical systems," *Journal of Machine Learning Research*, vol. 19, no. 29, pp. 1–44, 2018.

[10] V. Verdult, N. Bergboer, and M. Verhaegen, "Identification of fully parameterized linear and nonlinear state-space systems by projected gradient search," *IFAC Proceedings Volumes*, vol. 36, no. 16, pp. 711 –716, 2003, 13th IFAC Symposium on System Identification (SYSID 2003), Rotterdam, The Netherlands, 27-29 August, 2003.

[11] P. Cox and R. Tthy, "Alternative form of predictor based identification of lpv-ss models with innovation noise," in *2016 IEEE 55th Conference on Decision and Control (CDC)*, 2016, pp. 1223–1228.

[12] L. Ljung, "System identification," in *Signal Analysis and Prediction*, A. Procházka, J. Uhlíř, P. W. J. Rayner, and N. G. Kingsbury, Eds. Boston, MA: Birkhäuser Boston, 1998, pp. 163–173, ISBN: 978-1-4612-1768-8.

[13] J. Suykens and J. Vandewalle, *Nonlinear modeling: Advanced black-box techniques*. Springer US, 2012, ISBN: 9781461557036.

[14] G. Pillonetto, "The interplay between system identification and machine learning," Dec. 2016.

[15] T. Guo, Z. Xu, X. Yao, H. Chen, K. Aberer, and K. Funaya, "Robust online time series prediction with recurrent neural networks," in *2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, 2016, pp. 816–825.

[16] Yu Wang, "A new concept using lstm neural networks for dynamic system identification," in *2017 American Control Conference (ACC)*, 2017, pp. 5324–5329.

[17] J. Gonzalez and W. Yu, "Non-linear system modeling using lstm neural networks," *IFAC-PapersOnLine*, vol. 51, no. 13, pp. 485 –489, 2018, 2nd IFAC Conference on Modelling, Identification and Control of Nonlinear Systems MICNON 2018.

[18] S. C. DeMarco, G. Lazzi, Wentai Liu, J. D. Weiland, and M. S. Humayun, "Computed sar and thermal elevation in a 0.25-mm 2-d model of the human eye and head in response to an implanted retinal stimulator - part i: Models and methods," *IEEE Transactions on Antennas and Propagation*, vol. 51, no. 9, pp. 2274–2285, 2003.

[19] A. Lakhssassi, E. Kengne, and H. Semmaoui, "Modifed pennes' equation modelling bio-heat transfer in living tissues: Analytical and numerical analysis," *Natural Science*, vol. 02, Jan. 2010.

[20] L. Bedin and F. S. V. Bazn, "On the 2d bioheat equation with convective boundary conditions and its numerical realization via a highly accurate approach," *Applied Mathematics and Computation*, vol. 236, pp. 422 –436, 2014.

[21] S. Sadiqbatcha, H. Zhao, H. Amrouch, J. Henkel, and S. X. Tan, "Hot spot identification and system parameterized thermal modeling for multi-core processors through infrared thermal imaging," in *2019 Design, Automation Test in Europe Conference Exhibition (DATE)*, 2019, pp. 48–53.

[22] S. Pan and K. Duraisamy, "Long-time predictive modeling of nonlinear dynamical systems using neural networks," *Complexity*, vol. 2018, pp. 1–26, Dec. 2018.

[23] M. Optics, *The bio-heat equation*.

[24] J.-J. Greffet and C. Henkel, "Coherent thermal radiation," *Contemporary Physics*, vol. 48, no. 4, pp. 183–194, 2007. eprint: https://doi.org/10.1080/00107510701690380.

[25] L. Klein, H. Hamann, Y.-Y. Au, and S. Ingvarsson, "Coherence properties of infrared thermal emission from heated metallic nanowires," *Applied Physics Letters*, vol. 92, May 2008.

[26] S. Lin and T. Chou, "Numerical analysis of the pennes bioheat transfer equation on skin surface," in *2015 Third International Conference on Robot, Vision and Signal Processing (RVSP)*, 2015, pp. 71–74.

[27] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[28] A. M. Schäfer and H. G. Zimmermann, "Recurrent neural networks are universal approximators," in *Artificial Neural Networks – ICANN 2006*, S. D. Kollias, A. Stafylopatis, W. Duch, and E. Oja, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 632–640, ISBN: 978-3-540-38627-8.

[29] I. Houtzager, J. van Wingerden, and M. Verhaegen, "Recursive predictor-based subspace identification with application to the real-time closed-loop tracking of flutter," *IEEE Transactions on Control Systems Technology*, vol. 20, no. 4, pp. 934–949, 2012.

[30] A. Getis and A. Griffith, "Comparative spatial filtering in regression analysis," *Geographical Analysis*, vol. 34, pp. 130 –140, 2002.

[31] S. Ocampo and N. Rodriguez, "An introductory review of a structural var-x estimation and applications," *Revista Colombiana de Estadstica*, vol. 35, Jan. 2011.

[32] L. Ljung and T. Söderström, "Theory and practice of recursive identification," 1983.

[33] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37*, ser. ICML15, Lille, France: JMLR.org, 2015, 448456.

[34] R. Tth, "Modeling and identification of linear parameter-varying systems," in. Jan. 2010, vol. 403.